

ENVIRONMENT DIRECTORATE

**Joint Meeting of the Chemicals Committee and the Working Party on Chemicals,
Pesticides and Biotechnology**

**ADVANCING THE PRACTICE OF RISK ASSESSMENT WITH ALTERNATIVE TESTING
STRATEGIES: STATE OF THE SCIENCE FOR READ ACROSS AND RISK ASSESSMENT
GUIDANCE**

Peter KEARNS, Terumi MUNEKANE
Email: peter.kearns@oecd.org, terumi.munekane@oecd.org

JT03399812

Complete document available on OLIS in its original format

This document and any map included herein are without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area.

This document is the draft report on “Advancing the practice of risk assessment with alternative testing strategies: State of the science for read across and risk assessment guidance”. This project is part of the WPMN work on Risk Assessment and Regulatory Programmes. The objective of this project is to develop recommendations on how to advance the use of alternative testing strategies for nanomaterials in risk assessments.

This report was reviewed by the Working Party on Manufactured Nanomaterials and amended based on comments received.

ACTION REQUIRED: *The Joint Meeting is invited to declassify this document by 7 October 2016.*

TABLE OF CONTENTS

ABSTRACT	4
INTRODUCTION	5
STATE OF THE SCIENCE REPORT – “STATE OF THE SCIENCE FOR ALTERNATIVE TESTING STRATEGIES IN RISK ANALYSIS”	7
ALTERNATIVE TESTING STRATEGY CASE STUDY – ATS METHODS AND NANO-TiO ₂	11
WORKSHOP PROGRAMME	14
CROSS-CUTTING IDEAS AND RECOMMENDATIONS.....	16
Develop and foster data sharing systems and collaborations.....	16
Recommendations	16
Data mine existing databases and literature.....	16
Recommendations	17
Perform environmentally and biologically relevant testing.....	17
Recommendations	18
Ensure consistency between studies	19
Recommendations	19
RECOMMENDATIONS TO START NOW FOR 3-5 YEAR RESULTS.....	21
Determine how MNs are similar to, and differ from, conventional chemicals.....	21
Harmonise testing procedures and reporting	21
Perform occupational and environmental exposure monitoring.....	21
Develop appropriate MN groupings	22
Develop Adverse Outcome Pathway frameworks	22
Consider complex conditions.....	22
Advance MN safer-by-design principles	23
CONCLUSIONS	24
REFERENCES	25
ANNEX: AGENDA	30

Figures

Figure 1. Distribution of endpoints analysed.....	12
---	----

ABSTRACT

The Organisation for Economic Co-operation and Development (OECD) Working Party on Manufactured Nanomaterials (WPMN) Steering Group on Risk Assessment and Regulatory Programme (SG-AP) Project, “Advancing the practice of risk assessment with alternative testing strategies: State of the science for read across and risk assessment guidance” coordinated several efforts culminating with a workshop in 2014 to develop recommendations on how to advance the use of alternative testing strategies for nanomaterials in risk assessments. The project team surveyed the state of the science in alternative test strategies (ATS) from a “multiple models” perspective to show areas of common findings from differing approaches, areas of greatest uncertainty, and priorities for follow up in applied research toward risk management of manufactured nanomaterials (MNs). Experts from academia, industry, public interest groups, and government researched, analysed, and discussed how alternative models could be used to advance the risk analysis of MNs. The objectives were to identify how ATS could be used in a risk analysis context to inform human health, ecosystem health, and exposure data needs for MN in the near term and longer term, and research needs to support the development of these strategies in the near future. Efforts included organising and producing a symposium, a case study on alternative methods in safety testing, a state-of-the-science report and three white papers, each used as background materials for discussion in a September 2014 workshop that developed recommendations to advance knowledge and fill key gaps in understanding. Workshop deliberations revealed that ATS are now being used for screening, and that, in the near term, ATS could be developed for use in read-across or categorisation decision making within certain regulatory frameworks. Participants recognised that leadership is required from within the scientific community to address basic challenges, such as standardisation of materials, techniques and reporting, designing experiments relevant to realistic outcomes, as well as coordination and sharing of large-scale collaborations and data. The recommendations will aid the proper development and implementation of relevant ATS for MN testing that will expedite the ability to identify high-risk MNs, and lead to more rapid, cost-effective, and reliable MN safety testing for specific risk management decision contexts.

INTRODUCTION

1. The Organisation for Economic Co-operation and Development (OECD) Working Party on Manufactured Nanomaterials (WPMN) is engaged in several ongoing efforts relevant to the use of *in vitro* and other alternatives to traditional animal toxicology studies. In recent years, several workshops and expert meetings have led to the development of reports and recommendations for the use of specific methods for testing nanomaterials. Recent efforts include WPMN, which has conducted a preliminary review of Test Guidelines for their applicability to MN (OECD, 2009), and the European Scientific Committee on Consumer Safety (SCCS, 2012) providing guidance on safety assessment of nanomaterials in cosmetics. These reviews concluded that many of the basic toxicological principles in the existing OECD guidelines are applicable to MN testing, but in some aspects, nano-specific considerations must be made. For example, the OECD guidance documents developed for chemicals do not include adequate information on physico-chemical (p-chem) characterisation, measurement, delivery, tracking, and dosimetry and, because of this, a document specifically addressing MN on sample preparation and dosimetry was developed (OECD, 2012). This pilot project seeks to build on these efforts by considering needs and requirements for testing in a risk assessment context. Specifically, the pilot focuses on the current state of knowledge and research needs to advance the use of alternative testing strategies (ATS), as opposed to individual methods.

2. ATS, or strategies that reduce or replace the use of animal testing, have the potential to expedite the evaluation of new and existing substances by reducing the time and resources required to generate data compared to that of conventional tests. They promise to provide rapid screening and detailed mechanistic and cellular level toxicity information. Ultimately, data from ATS methodologies are expected to improve regulatory decision-making. A diversity of approaches are proposed, and there is an immediate need to inform the development of these strategies in such a way that they produce relevant, reliable, and useable data that can be used for risk assessment and policy development. To address these needs, the OECD WPMN Pilot Project “*Advancing the practice of risk assessment with alternative testing strategies: State of the science for read across and risk assessment guidance*” surveyed the state of the science in ATS from a “multiple models” perspective to show areas of common findings from differing approaches, areas of greatest uncertainty, and priorities for follow up in applied research toward risk management of MNs. A multiple models approach considers requirements to build an adequate level of evidence with ATS to overcome uncertainties associated with reliance on a single test or battery. This project stemmed from an ongoing collaboration between the Society of Risk Analysis (SRA) and the OECD, led by Lorraine Sheremeta (Alberta Ingenuity Labs, Canada), with support from Yasir Sultan (Environment Canada), Myriam Hill (Health Canada), Andy Atkinson (Health Canada), Carsten Kneuer (BfR, Germany), and coordinated by Jo Anne Shatkin (SRA Councillor, and Vireo Advisors, USA).

3. The main objectives of this pilot project were to:

- 1) Create a database of methods and ATS being used
 - a. Identify major users, which materials have been studied, the systems tested, and how well established they are for MN and distinguish existing from emerging methods
 - b. Compare findings across different users of methods (e.g. zebrafish assays)

- 2) Perform meta-analysis of physical and chemical properties and endpoints, seeking to extrapolate relationships to toxicology in the OECD WPMN dossiers.
 - a. Analyse the ability to use a suite of methods to inform comparative assessment of nanoscale material property data to effects
 - b. Characterise uncertainty associated with predictive relationships and propose strategies to address uncertainties
 - 3) Identify steps needed for these methods to be widely adopted.
4. The overall goal was to inform the development of guidance for ATS use in approaches for risk analysis through data gathering and analysis, discussion amongst contributors, and expert recommendations.
5. These objectives were met by the activities described here with numbers aligning with pilot project objectives. The pilot project commenced with an educational and interactive workshop at the SRA 2013 Annual Meeting in Baltimore, Maryland in December 2013, a half day presentation and discussion of ATS- related issues identified during a preliminary research stage, and served as the focal point for planning deliverables for the workshop. Several efforts followed, including an investigation into the state of the science for alternative testing strategies for MN and a report on ATS in risk analysis, development of a database as a case study with nano -TiO₂ on the current status of methods (with a focus on alternative testing methods, or ATMs) in safety assessment, and three white papers on the topics of human health, ecological, and exposure considerations regarding the use of ATS for MNs in preparation for an SRA -organised workshop held in Washington, D.C . in September 2014 (Annex). The workshop brought together 58 experts from diverse backgrounds to discuss the incorporation of ATS into risk assessment for MNs, and to assess the potential for a multiple models approach to using ATS for risk screening and their reliability for predicting human outcomes that can increase confidence, decrease uncertainty, and inform risk-based decision making. The workshop summary reports highlight the recommendations for next steps to advance the use of ATS for risk analysis.
6. The outputs of the pilot project included:
- Four manuscripts submitted for a special issue in the journal *Risk Analysis* , including the workshop report (Shatkin *et al.*, 2016), state-of-the-science report (Shatkin and Ong, 2016) and papers resulting from the white papers and workshop discussions of the human health (Stone *et al.*, 2016) and exposure groups (Sharma *et al.*, 2016).
 - A piece highlighting the policy aspects of the workshop was published on a popular web site, NanoWerk in advance of the workshop (JA Shatkin and L Sheremeta, 2014. Nanomaterial safety: An international collaboration on in vitro testing strategies; <http://www.nanowerk.com/spotlight/spotid=36452.php>)
 - Following the workshop, presentations were given to diverse audiences, including: the 2014 OECD Expert Workshop on Categorisation of Manufactured Nanomaterials (September 17-19, 2014, Washington, DC) , the Sustainable Nanotechnology Organization (November 2-4, 2014, Boston, MA), and a Symposium that included the paper authors at the Society for Risk Analysis Annual Meeting (December 7-10, 2014, Denver CO).

STATE OF THE SCIENCE REPORT – “STATE OF THE SCIENCE FOR ALTERNATIVE TESTING STRATEGIES IN RISK ANALYSIS”

7. Research into the available methods, key research groups and state of development led to the report “State of the Science for Alternative Testing Strategies in Risk Analysis.” The report reflects on the current status of alternative ATS for MNs and their application in screening level decision making, or prioritisation of substances for both human and ecological health. It captures the current state of development and application of ATS for MNs, and provides in-depth reviews of several strategies and their approaches, including the materials and methods used. The work is briefly summarised here.
8. The focus of the state -of-the-science analysis was to highlight the current efforts in a lternative testing methods (ATMs) to inform ATS for use towards efficiently advancing the fields of nanotoxicology and risk assessment. Several expert deliberations have concluded that while standalone ATM s may contribute to basic mechanistic or toxicity knowledge, they will not be sufficient for use in quantitative risk assessment; rather, a battery of ATMs will likely be used in a weight-of-evidence approach (*e.g.*, Nel *et al.*, 2013a). Strategically incorporating multiple ATMs into ATS will allow for an understanding of human and environmental behaviour and toxicity of MN across endpoints, receptors and material groups.
9. Incorporation of multiple single parameter *in vitro* tests in an ATS increases confidence in res ults, provides valuable information regarding potential mechanisms of effect, and identifies doses for further testing and aids in MN hazard and risk analysis. For instance, a combination of tests such as tetrazolium -based assays (*e.g.* MTT, MTS, XTT), trypan blue, alamar blue, lactate dehydrogenase assay, and neutral red uptake are regularly used to assess overt toxicity of MN (Nogueira *et al.*, 2014). Similarly, methods such as DCF fluorescence, lipid peroxidation, and assays measuring oxidative stress enzymes (*e.g.* glutathione, superoxide dismutase) are commonly employed to assess a MN’s potential to induce oxidative stress (Nogueira *et al.*, 2014). Furthermore, inflammatory response to MN exposure is often assessed by enzyme-linked immunosorbent assay (ELISA) that can be multiplexed to measure cytokines and chemokines (Nogueira *et al.*, 2014), and genotoxicity studies often employ the Comet assay, the micronucleus assay, the chromosome aberrations test, and the bacterial reverse mutation assay (Nogueira *et al.*, 2014; Magdolenova *et al.*, 2013). Research must continue to ensure that these tests are representative of *in vivo* eukaryotic conditions; for example, the OECD recently concluded that the commonly used Ames test, a bacterial mutagenicity assay, may not be suitable for detecting potential human genotoxicity induced by MNs because of the lack of endocytosis and limited MN diffusion across the bacterial cell wall (OECD, 2014a).
10. In addition to the ability to adopt these existing chemical testing methods for MN, there is concurrent development of emerging *in vitro* tests to address specific MN toxicity endpoints, such as reactive oxygen species generation. Single parameter tests have been used to develop high throughput screening (HTS) and high content screening (HCS) approaches that allow for the large number of unique MN formulations to be screened in a relatively rapid manner. HTS involve screening materials in batches, typically at rates of hundreds or thousands of readings per day and may take advantage of automated equipment, such as robotic liquid handling and/or computerised image capture. Examples include real-time characterisation of MN using dynamic light scattering (Wang *et al.*, 2013), cell-microelectronic sensing for cytotoxicity, (Moe, 2013) cell microarrays and Lab -on-a-Chip based screening of genotoxicity (Vecchio, 2014; Safe Work Australia, 2013) and zebrafish embryo studies determining effects on hatching, stress and development (Lin *et al.*, 2011). Large scale studies investigating small variations across MN have so far

contributed to understanding the correlations between physico-chemical properties and MN toxicity, and provide quantitative, mechanistic, pathway-based data (Nel *et al.*, 2013a). These examples of MN-based studies are useful for early tier hazard ranking, prioritisation for advanced testing, and determination of starting concentrations for further experiments (Lin *et al.*, 2011; Kavlock *et al.*, 2012).

11. *In vitro* models are becoming increasingly sophisticated and better at simulating human-relevant conditions. Three-dimensional (3D) cell co-cultures and (micro)fluidic models are emerging techniques used to create more realistic exposure conditions by simulating the morphology and physiology of natural tissue (Astashkina and Grainger, 2014; Roth and Singer, 2014; Rothen-Rutishauser *et al.*, 2005; Kostadinova *et al.*, 2013; Chortarea *et al.*, 2015; Horváth *et al.*, 2015). The results of these studies have shown that the 3D fluidic systems may more closely reflect *in vivo* conditions. While complex 3D systems may more closely resemble human conditions, these technologies are still in the early stages of development, and important information can also be rapidly and cost-efficiently obtained from combinations of less complex *in vitro* tests, especially when used as part of an ATS. A tiered ATS can be envisioned in terms of “increasing system complexity” where substances are first tested in the “simpler” more high-throughput *in vitro* systems (for example to assess overt toxicity) before deciding whether it is necessary to test that substance in a more complex *in vitro* system that more closely mimics the human situation.

12. *Ex vivo* studies are performed with tissues or organs collected from organisms while structure and viability are maintained in as normal condition as possible. In some cases, tissues or organs may be obtained from humans (*e.g.*, for use in OECD TG 428 *in vitro* Skin Absorption) or from animals sacrificed in slaughterhouses (*e.g.*, for use in OECD TG 437 Bovine Corneal Opacity and Permeability and OECD TG 438 Isolated Chicken Eye Test), excluding the need for animal sacrifice specifically for the experiment. *Ex vivo* studies can provide complex and realistic conditions and greater control over experimental parameters, while obtaining more results from the same number of organisms than in *in vivo* methods.

13. *Ex vivo* models may be used to evaluate MN penetration, uptake, and distribution, as well as toxicokinetics. For example, *ex vivo* rabbit and murine lung models (Beck-Broichsitter *et al.*, 2009; Nassimi *et al.*, 2009), bovine eye models (Kompella *et al.*, 2006) rat lower intestine model (Sandri *et al.*, 2010), immune responses in human blood cells (Delogu *et al.*, 2012; Mo *et al.*, 2008), cardiovascular effects (Stampfl *et al.*, 2011), and crossing of the placenta (Sønnegaard Poulsen *et al.*, 2013), have been employed. Comparison of *ex vivo* and *in vivo* gut exposure to TiO₂ MN show similar modes of translocation and localisation in the epithelium of both models (Brun *et al.*, 2014). Rat, mouse, or human precision cut lung slices are used to assess toxicant effects on the respiratory tract, such as cytotoxicity, genotoxicity, altered protein content, oxidative stress, apoptosis, cytokine release and histological changes (Wohlleben *et al.*, 2011; Hirn *et al.*, 2014; Kreyling *et al.*, 2014). The use of *ex vivo* methods for MN is still uncertain due to susceptibility to interferences, highly variable results, and an inability to determine appropriate dose-metrics that relate to *in vivo* exposure (Sauer *et al.*, 2014; Kim *et al.*, 2014). *Ex vivo* methods may be useful for prioritisation and ranking of MN toxicity (Kim *et al.*, 2014; Wohlleben *et al.*, 2011), but further method optimisation is likely needed to address reproducibility, false positives, false negatives, and to better understand the limitations of the system (Sauer *et al.*, 2014; Hirn *et al.*, 2014).

14. *In silico* methods (sometimes referred to as “non-testing methods”) are computational techniques that are crucial for the analysis of MN effects data. Traditional *in silico* methods have to be adapted, or new approaches must be developed for MN, due to the size-related properties of MN different from conventional chemical substances. The ultimate goal of *in silico* method development is to have a set of standard predictive models with defined parameters that can accurately and efficiently predict human and ecological toxicity of MN with minimal biological experimentation.

15. A lack of availability of quality data that can address the issues related to categorisation and grouping of MN based on their p-chem properties, mode of action or relevant exposure also hinders the development of *in silico* methods (Tantra *et al.*, 2014). Data from such approaches combined with clear reporting guidelines for MN studies can aid in increasing the predictability of *in silico* methods. These data may exist, but have not been reviewed or organised into a format to evaluate the replicability across substances.

16. It is generally accepted that no stand-alone *in vitro* or *ex vivo* test can replace a standardised *in vivo* method; however, a combination of such methods in a tiered strategy or integrated approaches to testing and assessment (IATA) will allow for prediction of potential relevant biological outcomes. Well-designed ATS will provide answers to focused and relevant MN toxicity questions. There are a number of different nano-specific ATS under development.

17. IATA can be used to identify and prioritise MN safety research needs, to assess the safety of a chemical using ATMs, and identify situations where *in vivo* testing is not needed. Generally, IATA consist of a tiered or semi-tiered framework with: 1) evaluation and organisation of existing data (using tools such as Adverse Outcome Pathways [AOPs]); 2) measurement of p-chem properties; 3) evaluation of the life cycle and biokinetics of the MN; 4) selection of appropriate context-specific toxicity tests (e.g. p-chem properties, use, release, potential exposure scenarios); and 5) application of a weight of evidence (WoE) analysis ('evidence based approach'), that considers and evaluates (based on the type and quality of data), all the results from the previous steps to reach a conclusion about the MN in question. IATA have been developed for skin irritation and corrosion (OECD, 2014b) human health risks of MN in food (Cockburn *et al.*, 2012), medical applications (Dusinska, 2013), and ecological assessments (Oomen *et al.*, 2014).

18. AOPs are conceptual frameworks being developed for risk assessment; they describe a sequential chain of causally linked events that lead to an adverse human or ecological outcome. Existing data can be harnessed to develop an AOP, which starts from a molecular initiating event, which links to key events at different levels of biological organisation (e.g., cellular or organ response), eventually leading to an adverse outcome at an organism or population level (Ankley *et al.*, 2010; OECD, 2013). It has become clear that direct correlations between a single MN p-chem property and *in vivo* outcomes are not possible; AOPs instead focus on groupings based on both the chemical activity and the consequent biological processes (OECD, 2013). Development of data using a number of different alternative testing methods for each key event builds WoE and confidence in these groupings. OECD has also published a guidance document on 'Developing and assessing adverse outcome pathways' (OECD, 2013). One major effort towards AOP development is the AOP Knowledge Base (<http://aopkb.org/>), an OECD online initiative that provides users guidance and tools to develop new AOPs, and allows users to contribute to the existing knowledge base, and to maximise the potential of already developed AOPs for risk assessment. As part of this initiative, OECD launched the user-friendly AOP wiki (<https://aopwiki.org>), designed to capture the scientific information in a central repository.

19. For screening MN of unknown toxicity, an approach in which all available qualitative and/or quantitative data are taken into account and weighted to help decide whether there is adequate information to support a decision, is important for risk assessment purposes. For example, a number of organisations have recommended measuring p-chem properties of MN using a multi-technique approach to overcome the limitations of individual methods to characterise MN in complex matrices (Methner *et al.*, 2009). A similar approach requiring multiple different assays to test one endpoint, will improve the utility of *in vitro* assays. Evaluation of the available data and ATM scores with respect to relevance and reliability is, however, a prerequisite for such WoE approaches.

20. In summary, ATMs are generally adaptable for MN testing, but nano-specific factors have to be taken into account. For example, the kinetics of MNs may be different from traditional chemicals, affecting

the behaviour in traditional media (*e.g.* settling to the bottom of a well). The use of realistic barrier models (*e.g.* 3-D co-cultures of dermal structures) appears to be an important consideration for MNs since MNs have different uptake mechanisms, and these types of models are likely more representative and correlated to *in vivo* outcomes. Several types of ATMs are described in the report – sections on *in vitro*, *ex vivo*, *in silico*, and high throughput screening (HTS) covered a spectrum of alternative tests. Emerging *in vitro* and *ex vivo* tests are currently being thoroughly tested by a number of groups such as the National Institute of Environmental Health Sciences' (NIEHS) Nano GO (Xia *et al.*, 2013), the German Federal Ministry of Education and Research's NanoCare and Cell@Nano (NanoCare 2009), and numerous academic labs. Some tests are being validated for their applicability to MNs by a number of agencies, including the OECD, the European Union Reference Library for alternatives to animal testing (EURL-ECVAM), and the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM). Novel *in silico* methods are steadily growing, and are an important part of MN toxicity testing. At this point, *in silico* models are not well enough developed, nor are there enough reliable data on MN toxicity to validate these models. HTS and high content screening (HCS) are upcoming methods that will allow for batch screening of MNs that will allow for higher testing volumes at a faster rate and more economical cost.

21. IATA were further elaborated in the human health section of the workshop by Dr. Vicki Stone. Strategic nano-testing approaches such as NanoTest's multiple models approach (Dusinska *et al.*, 2013), the NanoSafety Cluster Working Group 10's integrated approach (IATA) (Oomen *et al.*, 2014), a decision tree approach (Balls *et al.*, 2012), The University of California's Centre for the Environmental Implications of Nanotechnology's (UC CEIN) efficient and comprehensive HTS platform for predictive toxicological testing (George *et al.*, 2011; George *et al.*, 2010; Zhang *et al.*, 2012; Cohen *et al.*, 2012), a dynamic energy budget approach (Holden *et al.*, 2013; Klanjscek *et al.*, 2013; Muller *et al.*, 2014), an embryonic zebrafish metric (EZ metric) method (Liu *et al.*, 2013a; Liu *et al.*, 2013b), EPA's ToxCast screening approach, and Engineered NanoParticle Risk Assessment's (ENPRA) extensive *in vitro*, *in vivo*, and *in silico* testing programme (Winkler *et al.*, 2013; Dix *et al.*, 2006; ENPRA 2014). While all these strategies differed in approach, general themes emerged for their use in MN testing, including the inclusion of extensive p-chem characterisation, the use of multiple tests and models within a strategy, and the development of a tiered or semi-tiered scheme prior to the commencement of tests.

22. This report, in combination with the information presented and discussed, formed the basis of two manuscripts, the state-of-the-science paper, "Alternative Testing Strategies for Nanomaterials: State of the Science and Considerations for Risk Analysis" (Shatkin and Ong, 2016), and a workshop paper, "Advancing risk analysis for nanoscale materials: Report from an international workshop on the role of Alternative Testing Strategies" (Shatkin *et al.* 2016), both published in the peer reviewed journal *Risk Analysis*.

ALTERNATIVE TESTING STRATEGY CASE STUDY – ATS METHODS AND NANO-TiO₂

23. Project deliverables also included a case study in the form of a database and analysis. We sought to collect information about studies using *in vitro* and other alternative testing strategies, particularly also when *in vivo* testing was included, to examine the availability and quality of information. Nanoscale titanium dioxide (nano-TiO₂) was selected as a case study because, 1) it is a data-rich MN relative to others, and 2) it was included in the OECD WPMN Testing Programme on Manufactured Nanomaterials, 3) it is widely used and 4) it is considered “not soluble” and does not shed ions as do other metal oxides. In general, data inclusion criteria involved: 1) findings reported in the last 5 years; 2) findings for ATS involving several methods, rather than standalone assays; and 3) use of modified protocols for the study of engineered MNs. Studies that filled pre-identified knowledge gaps, such as oral exposure data, papers with high-throughput methods, and “round robin” inter-laboratory testing (where the same test was performed independently in many different labs), were among exceptions to the criteria above. From these, ninety-six studies and publications of *in vivo* and *in vitro* tests were identified and added in the database, including the OECD Working Party of Manufactured Nanomaterials (WPMN) Draft Dossier on Titanium Dioxide. Initially, a review of toxicology literature employing *in vitro* and other alternative methods and/or ATS with nano-TiO₂ was performed to identify potentially relevant data. We prepared a spreadsheet to capture the study details in a consistent searchable format. Thirty eight (38) variables for each study were identified for inclusion to capture key details of the study parameters and methods.

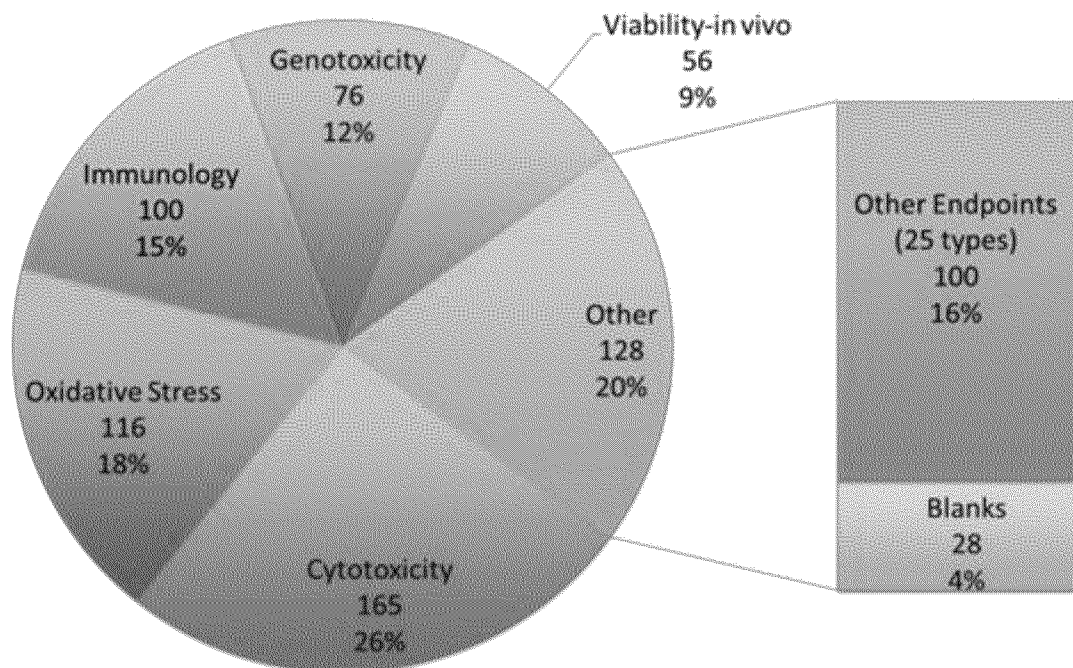
24. The data for the case study are taken from 96 unique publications on nano-TiO₂, and includes 1,820 “data entries”, and 38 categories, representing 69,160 data points. One data entry is representative of a result from an exposure to one type of nano-TiO₂ based on one type of endpoint measured by a single method (e.g. cytotoxicity by exposure to P25, as measured by the alamar blue assay) within a publication, thus one publication may warrant several data entries, particularly since larger comparative studies were prioritised. The 38 data categories include: discrete data (numerical and textual); pre-determined or refined categories, e.g., cytotoxicity as endpoint analysis type; and purely descriptive text, e.g. exposure notes. A few examples of the data recorded are: type of TiO₂ material, testing method, which organism/cell was used, doses used, doses at which effects were observed, etc.

25. An analysis of the data addressed 13 questions regarding the scope and magnitude of available ATS data in the case study. The analysis revealed that 64% of the studies were with commercially purchased materials, while 27% were synthesised by the researchers themselves. The top 5 types of endpoints studied were cytotoxicity (26%), oxidative stress (18%), immunology (15%), genotoxicity (12%), and *in vivo* viability (9%) (Figure 1). 16% had ‘other endpoints’ and in 4% of the studies no endpoint was defined (‘blanks’). The most commonly employed *in vitro* assays to study cytotoxicity were lactate dehydrogenase (LDH), MTT, WST-1, oxidative stress: DCF-DA, and DCFH-DA. For genotoxicity, the Comet assay was most common. Although data relevant to ecological exposures and the oral exposure route were prioritised, relatively few of such studies were available; the database contains more data relevant to inhalation exposure than from other exposure routes (such as dermal or ingestion). *In vivo* studies were performed mostly in rats, whereas *in vitro* experiments were mostly with human cells.

26. Examination of the database revealed gaps in the literature, and inconsistencies in performing, as well as in reporting assays in the experiments. We found that studies measured doses with a diversity of dose-metric (e.g. mg/L, µg/cm², etc.) regimes, preventing direct comparisons. Further, direct light exposure may increase the activity of nano-TiO₂ and change the biological effects during exposure, making this an

important factor in experimental design; however, our analysis revealed that many studies do not report lighting conditions, making it a challenge to determine the photoactive state of the particles.

Figure 1. Distribution of endpoints analysed



27. For example, *p*-chem properties such as the crystalline structure (*i.e.* ratio of anatase and rutile) and photoactivity, which can play a role in toxicity, are often not reported (Braydich-Stolle *et al.*, 2009; Ma *et al.*, 2012). Half of all results demonstrated no toxicity, while half reported effects at widely varying concentrations. These reported differential responses to nano-TiO₂ might be attributed to differences in the MN properties, cell types, dose and assay conditions, and lack of proper controls. We noted that a high proportion of reports did not account for MN -assay interferences, similar to previous observations noting that < 10% of peer-reviewed papers included appropriate controls for MN assay interference (Ong *et al.*, 2014). The observations made based on this case study highlighted the need for more standardised protocols for characterisation, testing and data reporting as well as guidance or criteria for the evaluation of study reports. Evaluation criteria might relate to the use of particular protocols and avoidance of interferences.

28. Importantly, we recorded the experimental doses tested, and found that only 3% (1 of 35) of aquatic *in vivo* studies used a relevant environmental dose of ≤ 16 $\mu\text{g/L}$ Predicted Environmental Concentration (highest emission scenario; Mueller and Nowack 2008), 0% (0 of 58) of studies used relevant inhalation doses within even orders of magnitude of the National Institute for Occupational Safety and Health (NIOSH) Recommended Exposure Limit of 0.3 mg/m^3 for nano-TiO₂ (NIOSH 2011) and 0% (0 of 8) used relevant ingestion doses of 3 $\mu\text{g/kg}$ body weight/day (Weir *et al.*, 2012), indicating a need for

environmentally- and human -relevant studies in addition to studies performed at higher doses for the purpose of hazard identification.

29. Many of the *in vivo* studies were performed at “overload” levels, where biological effects may be observed but do not allow for clearance of the MNs, so are therefore not representative of realistic exposures. These longer term exposure (chronic) studies were infrequently performed. Presumably, corresponding “chronic” ATM s are less accessible (or even not available) and more cost - and labour - intensive. There are significant gaps in ATS relevant to non -occupational exposure scenarios, including ingestion, dermal exposure, and non -mammalian organisms. Few studies addressed material transformations in biological matrices. The uncertainty associated with studies in both *in vitro* and *in vivo* systems under unrealistic exposure conditions limits the reliance on these data in a risk assessment context.

30. The case study was circulated to participants prior to the workshop, and presented as a poster at the workshop. A second phase of work, to link these data to a publicly available informatics database, was not completed, so the toxicity data are not currently linked to reported physical and chemical properties. Many of the data are from the WPMN Testing Programme dossier and could be readily linked at some future time. The overall goal is to make the database available to researchers and others. Utility of the database will demand that it be updated with a re -evaluation of studies for data quality and inclusion of phys-chem properties.

WORKSHOP PROGRAMME

31. The Society for Risk Analysis (SRA) organised a workshop on “Advancing risk analysis for nanomaterials: A workshop to explore how a multiple models approach can advance risk analysis of nanoscale materials” at George Washington University, Washington D.C. on September 15-16th, 2014. The programme can be found in Annex.

32. The workshop sought to survey the state of the science in alternative test strategies (ATS) from a “multiple models” perspective to show areas of common findings from differing approaches, areas of greatest uncertainty, and priorities for follow up in applied research toward risk management of nanomaterials (MNs). Specifically, the workshop was planned to consider how to incorporate ATS into risk assessment for manufactured MNs, and to assess the potential for a multiple models approach to using ATS for risk screening and their reliability for predicting human outcomes that can increase confidence, decrease uncertainty, and inform risk-based decision making.

33. The workshop highlighted shared strengths and gaps in support of a Weight of Evidence (WoE) approach relying on ATS to inform context-specific decisions about risk from exposure to novel nanoscale materials. The specific decision focus for this first “state-of-the-science and practice” evaluation was the initial set of decisions that a risk manager would need to make about seeking further (specific) data or declaring obvious safety for a novel nanoscale material. These decisions are typically termed screening - level decisions. The kinds of short -term alternative testing methods (ATMs) considered include *in vitro* and cell free assays that can be included in high throughput assays or the initial stages of tiered testing cascades.

34. To achieve the objectives established for the workshop, 58 participants including several from the WPMN and from academia, government, industry, and NGOs (see Addendum to this report [ENV/JM/MONO(2016)63/ADD] for full participant list) gathered for one and one half days. Three white papers were prepared by invited speakers on the topics of human health, ecological, and exposure considerations regarding the use of ATS for MNs, led by Professor Vicki Stone of Heriot Watt University, Dr. Patricia Holden of UC Santa Barbara, and Dr. Monita Sharma of PETA International Science Collaborative, respectively.

35. Outputs from the workshop include:

- 1) new collaborations through funding partners and collaborating organisations;
- 2) five manuscripts submitted for a special edition in the journal *Risk Analysis*, including the workshop report, papers resulting from the discussion in the human health, ecological health, and exposure groups;
- 3) web publications (JA Shatkin and L Sheremeta, 2014. Nanomaterial safety: An international collaboration on *in vitro* testing strategies

(<http://www.nanowerk.com/spotlight/spotid=36452.php>);

- 4) presentations to diverse audiences, including: the 2014 OECD Expert Workshop on Categorisation of Manufactured Nanomaterials (September 17-19, 2014, Washington, DC); Sustainable Nanotechnology Organization (November 2-4, 2014, Boston, MA); and a Symposium at the Society for Risk Analysis Annual Meeting (December 5-8, 2014, Denver, CO); and,
- 5) presentations and this report to OECD WPMN on the ATS Pilot Project, which includes workshop findings.

36. The workshop experts from diverse fields also included 40% of participants from outside the U.S., 28% from private organizations/industry, 33% from governmental agencies, 12% from NGOs and 27% from academia. Women made up more than half of the participants, and there were both early career and emeritus attendees.

37. Workshop attendees were charged with:

- 1) *How can the findings from ATS and ATMs be used in combination with conventional testing methods to reduce uncertainty and better inform screening-level risk assessments (with respect to human health, ecosystem health, and exposure)?*
 - a. *How can these findings **be used now** to amass a weight of evidence approach that supports risk assessment?*
 - b. *What strategies can be developed to increase the value of, and confidence in, the use of ATS findings in risk assessment?*
- 2) *What additional work is needed in the near-term (3-5 years) so that ATS methods can better support risk assessments that inform screening-level risk management decisions (with respect to human health, ecosystem health, and exposure)?*

38. The workshop was structured into 3 sessions for each topic and a poster session, starting with all participants being involved in the state-of-the-science (SOS) presentations and group roundtables, then the members divided into breakout groups (full programme found in Annex). For the SOS presentations, three white papers were developed and presented for each of these topics by leading experts in each field, followed by 3-person panel list discussions, and finishing with interactive audience participation. These presentations helped formulate a common foundation amongst participants for the breakout sessions. Based on these presentations, workshop participants held roundtable discussions as a group, during which these topics were further elaborated and discussed from different viewpoints. Then, participants joined one of three breakout groups (human health, ecological health, and exposure assessment issues) where focused discussions were held regarding the concepts in the white papers, the presentations, and recommendations for research needs and next steps. Each breakout group presented IATA conclusions to all participants, and the workshop finished with a plenary discussion, conclusion, and summarising of the next steps.

39. Below is a summary of the conclusions followed by recommendations.

CROSS-CUTTING IDEAS AND RECOMMENDATIONS

40. Experts and participants agreed that ATS are useful as they, for the most part, require fewer resources, are less time intensive, are generally higher throughput, and are cost effective. A multiple models approach incorporating IATA, tiered approaches, and WoE analysis will increase confidence but require effort to coordinate and standardise testing and methods. At this stage, the results from these tests will provide supplemental value to leverage better and more relevant data from whole -animal tests and allow for better interpretation of the results. ATS can be useful for screening, prioritisation, and reduction of uncertainty, as long as potential limitations, especially those related to MN (e.g., lack of physiological complexity, issues of reproducibility, metrology, distinction between acute and chronic exposures and effects, and potential for interferences) are considered. It is envisioned that these data will allow for the development of more focused/targeted tests that will inform risk assessment for next 3 -5 years. Key recommendations from the breakout groups overlapped significantly and are summarised here:

Develop and foster data sharing systems and collaborations

41. Development of ATS relies on contributions from many different areas of the nanoscience community. Risk assessment of MN would benefit from improved access to large data sources, enabling more efficient collaboration. The development of open databases, shared forums, or other data -sharing formats capable of storing large amounts of data could strengthen the ability to perform risk assessment and to share knowledge. Participants indicated that although a number of large databases exist that contain information on useful assays, targets of accumulation, and *in vitro* and *in vivo* effects, many of these databases are not easily accessible. Publication bias against negative results can present an unbalanced view of MN toxicity; to counter this, researchers should be conscious of this bias when performing hypothesis-driven investigations, editors must be willing to accept negative results, and risk assessors must take care when using published literature.

Recommendations

- Specify standards for the funding of studies or publication of results, especially for funders and publishers, who can determine requirements such as availability, consistency and quality of data;
- Create platforms for data sharing and storing large amounts of data (e.g. user-friendly databases, wiki, *etc.*); and
- Incentivise researchers to make data public.

Data mine existing databases and literature

42. A wealth of relevant data exists in the published literature, OECD WPMN Testing Programme dossiers, industrial datasets, and other such compilations. This information can be used to assess testing protocols, substantiate results, and determine replicability of studies. Including data mining as part of an integrated testing system can help screen and identify MN, pathways, endpoints, and data gaps of particular concern.

43. Research gaps in the environmental and occupational data, and the nano-TiO₂ case study developed for this workshop identified a low level of consistency in data collection and reporting, and a diversity of reported responses, which limit the ability to compare studies for specific toxicological outcomes. Therefore, data mining and comparison of results may be improved by creating datasets of well-characterised materials in carefully designed and/or standardised studies. Data mining efforts, such as analysis of p-chem properties, in the context of relevant nano-specific characteristics, as part of categorisation and grouping of MN, will advance the understanding of key testing parameters and relationships.

44. In data mining, it is essential to incorporate some element of data quality evaluation of previously developed information, in terms of certainty around materials tested (such as consistent terminology and detailed p-chem characterisation) and consistency in data obtained. Some participants suggested using a publication date (e.g., after 2010) as a threshold for acceptance of studies, while others suggested categories for data quality (e.g. the study reports a minimum set of p-chem measurements).

Recommendations

- Develop and implement robust statistical methods appropriate for mining the existing data sets;
- Determine, based on representativeness and validity, which ATMs are appropriate for use and correlate to human or environmental health, especially low dose chronic exposure;
- Ensure that researchers consistently report relevant experimental details, such as context- and receptor-specific parameters (e.g., relevant exposure route);
- Use these data to refine both ATM s and *in vivo* testing to avoid redundancy and wasted resources; and
- Develop protocols and criteria for assessing data quality of all existing data for screening information.

Perform environmentally and biologically relevant testing

45. All three breakout groups agreed on the potential for using ATS for testing MN at relevant concentrations, under more realistic exposure conditions and for screening relative to traditional substances and toxicological effects of bulk materials. Deliberation, determination, and assessment of potential realistic situations, such as the conditions of MN release and subsequent MN life-cycle transformations will help to direct ATS development.

46. Using relevant dosimetry based on findings from occupational or environmental monitoring studies will be helpful in assessing risk based on relevant route of exposures. Although high concentration, short duration exposures may be informative to determine concentrations that can elicit toxicity, the mechanisms of toxicity, and to elucidate health effects associated with incidents that result in high levels of acute exposure (e.g., from spills or other accidental releases), repeated and/or chronic exposure at lower concentrations is generally more realistic for occupational, consumer and environmental exposures. However, there are currently no validated *in vitro* tests to mimic chronic, repeated exposures; development and validation of these tests are an immediate need for MNs. At a minimum, standardising test conditions to represent meaningful exposure modes, and concentrations are implementable now and are a practical complement to short duration, high concentration exposures for predicting MN effects.

47. One issue emphasised during the discussions is the need for reliable analytical methods to measure MN in complex biological and environmental media to aid context-specific testing. There are established tools and techniques to characterise MN in their pristine form, but most of the techniques have limitations when it comes to assessing MN in complex ecological and biological matrices (von der Kammer *et al.*, 2012). Such limitations could be overcome by developing new tools, by modifying existing techniques, and by using multiple techniques to characterise MN. Despite ongoing efforts by WPMN, ISO and others to develop a standardised list of essential p-chem parameters, this remains a challenge.

48. Research determining the fate of MN in both human and ecological exposure scenarios will help identify the target organisms and organs most susceptible to MN exposure. Identifying the route of exposure of a MN will highlight the organs, systems, or organisms likely to receive the highest initial dose, then toxicokinetic studies and/or appropriate physiologically based pharmacokinetic (PBPK) models, can direct researchers to focus on cells, organs, and tissues that may be in contact with MN. The biologically effective dose, or the amount of contaminant that interacts with the internal target tissue or organ, may be different than the initial measured or administered dose of MN. In addition, identification of at-risk and/or sensitive populations can help prioritise testing.

49. Further, biotic and abiotic molecules, such as serum albumin or natural organic matter, can bind and interact with MN and change their surface p-chem properties and distribution; these molecules can form a corona around the MN, altering bioavailability, transport, toxicity, etc. These alterations will change how the particles are “seen” by the receptor, at the biological exposure sites. For example, MN incorporated into food products could undergo transformation in the saliva, which could in turn change how they are presented to gut microflora. In aquatic systems, MN may interact with natural organic matter, such that it will not be in pristine form when taken up by detritivores. Protocols can be developed to ensure that ATS are testing the appropriate MN form, and that they therefore represent the most realistic exposure scenario. For example, MN of respiratory concern could be first coated in natural lung lining fluid before presentation to cells, and MN of aquatic concern could be incubated in natural waters before being introduced to aquatic fauna. Biologically relevant exposures, such as incubations in natural waters are important, but exposures in other media may be necessary to determine if the effects seen, or not seen can be attributed to the biologically and environmentally relevant media.

50. Simulating realistic scenarios is of utmost concern to improve risk assessment and to determine which areas are of highest concern; in particular, it will be necessary to focus on environmental realism to help illuminate which impacts beyond acute toxicity (notably lethality) may be of concern (e.g., by incorporating variables to assess low-dose chronic toxicity, nutrient cycling, UV intensity). Until our understanding of MN behaviour is improved, comprehensive p-chem measurements under these different exposure scenarios will be beneficial for inferring MN toxicity, at least for screening analyses. Due to transformations that take place during a MN life-cycle (e.g., aggregation/agglomeration, dissolution, corona formation), MN p-chem characterisation is necessary at various stages throughout the product life-cycle (i.e., as manufactured, as tested, as used, and as released following disposal) to understand how modification of p-chem characteristics can influence the MN behaviour. The ultimate goal is to understand the connections between p-chem properties, exposure, hazard, and toxicity, so as to minimise the need to fully characterise and test every material for risk assessment.

Recommendations

- Identify routes of exposure of MNs and use relevant models;
- Use context-specific doses;
- Measure at relevant time points;

- Move forward with toxicokinetics testing, as a pre-requisite for PBPK modelling;
- Use appropriate biological or environmental media to ensure MN has formed realistic corona;
- Measure MN p-chem changes throughout the material life-cycle;
- Develop tools, instrumentation, and/or assays to measure MN p-chem characteristics;
- Develop a standardised list of essential p-chem characteristics;
- Test worst-case conditions;
- Identify at-risk and sensitive populations;
- Incorporate biological and environmental complexity;
- Test the most sensitive groups or species.

Ensure consistency between studies

51. Determination of relevant dose and standardisation of dose metrics has long been a challenge for MN, and has hindered comparison between *in vitro* and *in vivo* studies. The nano-TiO₂ case study found that *in vitro* experiments relevant to inhalation exposure often used mass per surface area and mass per volume, however mass per animal weight, or mass per unit volume of air is most often used for *in vivo* dosing. If there were a standard set of p-chem reporting requirements, (e.g., shape, density, mass, and concentration), then surface area and particle number could also be calculated.

52. Research to determine suitable positive and negative material controls or reference materials is still ongoing; gathering more information about MN p-chem properties and using well-characterised reference materials (e.g., from National Institute of Standards and Technology [NIST]) or representative industrial nanomaterials (e.g., from Joint Research Centre [JRC] repository in Ispra) for comparison purposes would facilitate progress toward this goal. The inclusion of conventional or ionic controls (in the case of soluble MNs) will also be important for such comparisons.

53. Many ATS include biochemical assays and other established methods used for conventional chemicals, but MN can potentially interfere and generate inaccurate results possibly leading to the determination of false positives or negatives, and/or to the creation of conflicting results among studies. For instance, MN have interfered with the optical density readings for tetrazolium-based assays such as MTS and MTT; however, removal of MN via centrifugation before reading the assay reduced the variations in generated data (Xia *et al.*, 2013). Some assays may be less susceptible to such effects, and interference should be assessed on a case-by-case basis (Kroll *et al.*, 2011, Ong *et al.*, 2014). Therefore, the use of a multiple models strategy, comparing results of different assays with the same intended outcome (e.g., MTT and alamar blue for cytotoxicity) across a number of studies can provide evidence for choosing appropriate assays. Furthermore, choosing multiple assays to evaluate the same toxicity endpoint can also help highlight and mitigate the limitations of any particular assay.

Recommendations

- Use comparable dose metrics between *in vitro* and *in vivo* studies;

- Control for MN-assay interference by removing MN, and/or using relevant negative and positive controls;
- Perform multiple assays that test for the same outcomes to increase confidence in results.

RECOMMENDATIONS TO START NOW FOR 3-5 YEAR RESULTS

54. While much of the work can be done immediately, the participants discussed which actions were needed now, or in the near future, to advance knowledge and practice in 3–5 years time. The following were suggested:

Determine how MNs are similar to, and differ from, conventional chemicals

55. Some MNs have similar toxicity profiles to their bulk or ionic counterparts, and therefore may not require additional testing. Benchmarking MNs against conventional chemicals could expedite testing and help prioritisation. To achieve this, diagnostics must be developed to determine when MNs behave more like conventional chemicals (e.g. those that dissolve completely into metal ions) and when their nanoscale properties create novel behaviour. Work performed on fine or particulate matter may be applicable for MNs. Models used for other chemicals, such as pesticides, may be applicable to MNs. Further route-specific similarities should be identified.

Harmonise testing procedures and reporting

56. Harmonisation of standardised operation procedures (SOPs), including reference materials and appropriate controls, media and conditions, and technologies/equipment, as well as harmonisation of information reporting will result in faster, more consistent, and more reliable data generation. However, until appropriate assays and tests are substantiated and standardised for MN, and the relevant MN properties are identified, coordination will be difficult. Ideally, SOPs will be applicable to a wide range of MNs, however it is possible that harmonisation may not be feasible due to the variability amongst MNs (e.g. metal ion shedding vs non-metal ion shedding MNs). In this case, the goal may be to produce SOPs for particular groups or categories of MNs, or include preliminary testing steps for appropriateness (e.g. do a wavelength scan of the MN itself prior to performing assays dependent on absorbance or fluorescence).

57. To ultimately achieve this goal, it would be practical to begin narrowing the list of available methods and excluding those not appropriate for MNs. For example, a commonly used cytotoxicity test, the MTT assay, is interfered with by many MNs, particularly at higher concentrations (Worle-Knirsh *et al.*, 2009; Holder *et al.*, 2012; Kroll *et al.*, 2012). The mechanism of interference differs between MNs and is difficult to predict, therefore its suitability for MN testing may be limited, especially in light of the numerous other ATMs available for cell viability testing. Methods and criteria to substantiate and validate ATMs need to be developed.

Perform occupational and environmental exposure monitoring

58. As more MNs are produced and incorporated into consumer products and services, exposure monitoring will be important for tracking of MN movement and concentration, and to allow for development and confirmation of *in silico* modelling. Therefore, sample collection, monitoring, and detection and measurement technologies for MNs in complex matrices, such as air, water, and soil need to be developed and improved to be sensitive enough to detect small particles at small concentrations, as well as for identification of materials. This will allow relevant exposure parameters such as dosimetry, conditions, and duration for ATS, and validate computational models for prediction purposes. In conjunction with this, development of *in silico* and informatics techniques should be prioritised to

minimise reliance on *in vivo* testing and limit our need to perform extensive, costly, and time-consuming testing.

Develop appropriate MN groupings

59. Grouping and read-across are concepts that may be used for MN risk assessment, but methods are currently inconsistent. A number of different methods, such as heat and self-organising maps to group MNs by bioactivity, the use of *in vitro*-chem data to inform structure-property relationships (SPRs), structure-activity relationships (SARs), and quantitative structure-activity relationships (QSARs), or Principal Component Analysis (PCA) methods to help identify the weight of contributions of different *in vitro*-chem properties to toxicity, can be employed. Existing data can be mined to establish rationale for groupings. This strategy may also be a quick method to prioritise MNs for further assessment, and can be used in many ways to get general overview or specific concerns. If groupings for MNs can be established, then tiered approaches and procedures can be developed for these groups. Furthermore, targeted testing can be performed as a way to increase confidence in the group identification.

Develop Adverse Outcome Pathway frameworks

60. ATS that incorporate adverse outcome pathway (AOP) frameworks can facilitate the connection between mechanistic and ecologically relevant / health related outcomes. AOPs provide a practical solution to data organisation and generation of meaningful information from the available data. AOPs build from a conceptual framework that describes a sequential chain of causally linked events at different levels of biological organisation that lead to an adverse biological or eco-toxicological effect. Similarly, the concept of exposure pathways can also be incorporated into AOP use in order to capture transformations of particles in key scenarios with underrepresented endpoints. AOPs are increasingly becoming more useful for risk assessment and reducing *in vivo* testing.

61. The continued gathering of mechanism-related data through *in vitro*, HTS, HCS, and other alternative methods will help identify nano-specific effects at each level. The improvement and consistency of data collection will aid cross-comparison and collection of data to be used in an AOP. Development of AOPs for MNs can be cultivated using the AOP-Wiki, a central repository for AOPs that serves as a platform for sharing AOP-related knowledge based on existing data. AOPs provide a framework for organising data from MN studies and can be used to develop intelligent testing strategies or IATAs based on mechanism.

Consider complex conditions

62. Currently, testing is usually performed in relatively straightforward conditions with pristine materials to allow for basic toxicity testing in single organisms or cells. While these studies improve our understanding of the underlying mechanisms of toxicity, as previously mentioned, these ultimately will not be representative of realistic situations. As ATMs and ATS are developed, we must consider the form of MNs in the final product, their entire life-cycle, and also consider co-exposures with other materials. More information is needed on low dose, chronic endpoints; these can be informed by acute studies, but must be performed, otherwise important effects will be missed. In the near term, we need to encourage risk-contextualised relevance in basic research, and create incentives to do research that is relevant to testing needs for decision-making. In order to fulfil this goal, risk exposure pathways and MN fate studies need to be performed, and assays with long-term relevance need to be developed and substantiated as appropriate for MNs.

Advance MN safer-by-design principles

63. Rather than perform safety testing post-development and production of a nano-enabled product or service, it is preferable to develop MN applications that are “safer-by-design”. To allow for this, recommendations and ATS can start to be compiled that can be used by industry to develop safe (or at least safer) nano-enabled products. Some materials can be ruled out, or prioritised, based on the results from alternative testing. Producers can start incorporating decision making feedback into material design, and should take a life-cycle perspective, incorporating current knowledge of how MNs behave in various matrices.

CONCLUSIONS

64. The OECD W PMN project successfully achieved its objectives of informing the development of guidance for alternative testing methods (ATMs) and alternative test strategies (ATS). As described here and in the Annex, the objectives were met. The case study and state-of-the-science report identified knowledge gaps in the research and data, inconsistencies amongst the testing methodologies and reporting of MN toxicity studies, and highlighted existing and emerging ATMs and ATS. The papers, presentations, and expert discussion stemming from the workshops further discussed these issues, highlighting the utility of ATS for screening level decisions at present. Ultimately, the efforts of the project produced a list of immediately implementable recommendations on how to improve the progress and growth of ATS using a multiple models approach.

65. The project generated specific and community-wide action items for advancing the use of ATS in MN risk assessment. The development, improvement, validation, and standardisation of realistic and relevant ATMs and ATS are critical to generating data that will support categorisation and read-across. The ongoing development of ATS will allow for screening and prioritisation of hazardous MNs, with the goal of eventually gaining enough information and understanding to move away from *in vivo* testing. As discussed, some unique considerations are necessary when adapting or developing ATS for MNs, but can be addressed through community collaboration and data sharing, performance of appropriate and applicable tests, and thorough reporting and dissemination of results. As we simultaneously continue to learn more about the performance of testing methods, the key predictive characteristics of MNs and their testing environments, and the realistic use of MNs and human and environmental exposure to MNs, salient risk assessments will require ongoing and iterative communication of findings across the diverse array of disciplines and sectors represented in this project.

REFERENCES

- Ankley, G.T. et al., 2010. Adverse outcome pathways: A conceptual framework to support ecotoxicology research and risk assessment. *Environmental Toxicology and Chemistry*, 29(3), 730–741.
- Astashkina, A. and Grainger, D.W., 2014. Critical analysis of 3-D organoid *in vitro* cell culture models for high-throughput drug candidate toxicity assessments. *Advanced Drug Delivery Reviews*, 69-70, 1–18.
- Beck-Broichsitter, M. et al., 2009. Pulmonary drug delivery with aerosolizable nanoparticles in an *ex vivo* lung model. *International Journal of Pharmaceutics*, 367(1-2), 169–178.
- Braydich-Stolle, L.K. et al., 2009. Crystal structure mediates mode of cell death in TiO₂ nanotoxicity. *Journal of Nanoparticle Research*, 11(6), 1361–1374.
- Brun, E. et al., 2014. Titanium dioxide nanoparticle impact and translocation through *ex vivo*, *in vivo* and *in vitro* gut epithelia. *Particle and Fibre Toxicology*, 11, 1–16.
- Chortarea, S. et al., 2015. Repeated exposure to carbon nanotube-based aerosols does not affect the functional properties of a 3D human epithelial airway model. *Nanotoxicology*, doi: 10.3109/17435390.2014.993344
- Cockburn, A. et al., 2012. Approaches to the safety assessment of engineered nanomaterials (ENM) in food. *Food and Chemical Toxicology*, 50(6), 2224–2242.
- Cohen, Y. et al., 2013. *In silico* analysis of nanomaterials hazard and risk. *Accounts of Chemical Research*, 46(3), 802-812.
- Delogu, L.G. et al., 2012. *Ex vivo* impact of functionalized carbon nanotubes on human immune cells. *Nanomedicine*, 7(2), 231–243.
- Dix, D.J. et al., 2006. The ToxCast Program for prioritizing toxicity testing of environmental chemicals. *Toxicological Sciences*, 95(1), 5–12.
- Dusinska, M. and NanoTEST consortium, 2013. Final Report: Development of methodology for alternative testing strategies for the assessment of the toxicological profile of nanoparticles used in medical diagnostics. Norway.
- ENPRA, 2014. Risk assessment of engineered nanoparticles: Final report.
- George, S. et al., 2011. Use of a high-throughput screening approach coupled with *in vivo* zebrafish embryo screening to develop hazard ranking for engineered nanomaterials. *ACS Nano*, 5(3), 1805–1817.
- George, S. et al., 2010. Use of a rapid cytotoxicity screening approach to engineer a safer zinc oxide nanoparticle through iron doping. *ACS Nano*, 4(1), 15–29.

- Hirn, S. et al., 2014. Pro-inflammatory and cytotoxic response to nanoparticles in precision-cut lung slices. *Beilstein Journal of Nanotechnology*, 5, 2440–2449.
- Holden, P.A. et al., 2013. Ecological nanotoxicology: Integrating nanomaterial hazard considerations across the subcellular, population, community, and ecosystems levels. *Accounts of Chemical Research*, 46(3), 813–822.
- Holder, A.L. et al., 2012. Particle-induced artifacts in the MTT and LDH viability assays. *Chemical Research in Toxicology*, 25, 1885-1892.
- Horváth, L. et al., 2015. Engineering an *in vitro* air-blood barrier by 3D bioprinting. *Scientific Reports*, 5, 7974.
- Kammer von der, F. et al., 2012. Analysis of engineered nanomaterials in complex matrices (environment and biota): General consideration and conceptual case studies. *Environmental Toxicology and Chemistry*, 31(1), 32-49.
- Kavlock, R. et al., 2012. Update on EPA’s ToxCast program: Providing high throughput decision support tools for chemical risk management. *Chemical Research in Toxicology*, 25(7), 1287–1302.
- Kim, Y.H. et al., 2014. Comparative lung toxicity of engineered nanomaterials utilizing *in vitro*, *ex vivo* and *in vivo* approaches. *Journal of Nanobiotechnology*, 12, 1-12.
- Klanjscek, T. et al., 2013. Dynamic energy budget approach to modeling mechanisms of CdSe quantum dot toxicity. *Ecotoxicology*, 22(2), 319–330
- Kompella, U. et al., 2006. Luteinizing hormone-releasing hormone agonist and transferrin functionalizations enhance nanoparticle delivery in a novel bovine *ex vivo* eye model. *Molecular Vision*, 12, 1185–1198.
- Kostadinova, R. et al., 2013. A long-term three dimensional liver co-culture system for improved prediction of clinically relevant drug-induced hepatotoxicity. *Toxicology and Applied Pharmacology*, 268(1), 1–16.
- Kreyling, W.G. et al., 2014. *In vitro* and *in vivo* interactions of selected nanoparticles with rodent serum proteins and their consequences in biokinetics. *Beilstein Journal of Nanotechnology*, 5, 1699–1711.
- Kroll A, et al., 2011. Cytotoxicity screening of 23 engineered nanomaterials using a test matrix of ten cell lines and three different assays. *Particle and Fibre Toxicology*, 8(1), 9.
- Kroll, A., et al., 2012. Interference of engineered nanoparticles with *in vitro* toxicity assays. *Archives of Toxicology*, 86(7), 1123–1136.
- Lin, S. et al., 2011. High content screening in zebrafish speeds up hazard ranking of transition metal oxide nanoparticles. *ACS Nano*, 5(9), 7284–7295.
- Liu, X. et al., 2013a. NEIMiner: nanomaterial environmental impact data miner. *International Journal of Nanomedicine*, 8(Suppl 1), 15-29.
- Liu, X. et al., 2013b. Predictive modeling of nanomaterial biological effects. *International Journal of Nanomedicine*, 8(Suppl 1), 31-43.

- Ma, H., Brennan, A. and Diamond, S.A., 2012. Photocatalytic reactive oxygen species production and phototoxicity of titanium dioxide nanoparticles are dependent on the solar ultraviolet radiation spectrum. *Environmental Toxicology and Chemistry*, 31(9), 2099–2107.
- Magdolenova, Z. et al., 2013. Mechanisms of genotoxicity. A review of *in vitro* and *in vivo* studies with engineered nanoparticles. *Nanotoxicology*, 8(3), 1–46.
- Malkiewicz, K., et al., 2011. Nanomaterials in REACH – evaluation of applicability of existing procedures for chemical safety assessment to nanomaterials. Project report.
http://www.steptoe.com/assets/htmldocuments/SKEPP%202011%20Nanomaterials_in_REACH_report_15082011.pdf
- Methner, M., Hodson, L. and Geraci, C., 2009. Nanoparticle emission assessment technique (NEAT) for the identification and measurement of potential inhalation exposure to engineered nanomaterials — Part A. *Journal of Occupational and Environmental Hygiene*, 7(3), 127–132.
- Mo, Y. et al., 2008. Cytokine and NO release from peripheral blood neutrophils after exposure to metal nanoparticles: *in vitro* and *ex vivo* studies. *Nanotoxicology*, 2(2), 79–87.
- Moe, B., Gabos, S. and Li, X.-F., 2013. Real-time cell-microelectronic sensing of nanoparticle-induced cytotoxic effects. *Analytica Chimica Acta*, 789, 83–90.
- Mueller, N.C. and Nowack, B., 2008. Exposure modeling of engineered nanoparticles in the environment. *Environmental Science and Technology*, 42(12), 4447–4453.
- Muller, E.B. et al., 2014. Impact of engineered zinc oxide nanoparticles on the energy budgets of *Mytilus galloprovincialis*. *Journal of Sea Research*, 94, 29–36.
- NanoCare, 2009. Health related aspects of nanomaterials: Final scientific report. Eds: Kuhlbusch, T.A.J., Krug, H.F., Nau, J.
http://www.nanopartikel.info/files/projekte/NanoCare/NanoCare_Final_Report.pdf
- Nassimi, M. et al., 2009. Low cytotoxicity of solid lipid nanoparticles in *in vitro* and *ex vivo* lung models. *Inhalation Toxicology*, 21(s1), 104–109.
- Nel, A., Xia, T., et al., 2013a. Nanomaterial toxicity testing in the 21st century: Use of a predictive toxicological approach and high-throughput screening. *Accounts of Chemical Research*, 46(3), 607–621.
- Nel, A.E. et al., 2013b. A multi-stakeholder perspective on the use of alternative test strategies for nanomaterial safety assessment. *ACS Nano*, 7(8), 6422–6433.
- Nogueira, D. et al., 2014. Mechanisms underlying cytotoxicity induced by engineered nanomaterials: A review of *in vitro* studies. *Nanomaterials*, 4(2), 454–484.
- NIOSH, 2011. Current Intelligence Bulletin 63 -Occupational Exposure to Titanium Dioxide. U.S. DHHS (NIOSH) Publication No. 2011–160.
- OECD, 2009. Preliminary review of OECD test guidelines for their applicability to manufactured nanomaterials. OECD Environment, Health and Safety Publications Series on the Safety of Manufactured Nanomaterials. No. 15. [ENV/JM/MONO\(2009\)21](#), Paris, France.

- OECD, 2012. Guidance on sample preparation and dosimetry for the safety testing of manufactured nanomaterials. OECD Environment, Health and Safety Publications Series on the Safety of Manufactured Nanomaterials. No. 36. [ENV/JM/MONO\(2012\)40](#), Paris, France.
- OECD, 2013. Guidance document on developing and assessing adverse outcome pathways. OECD Environment, Health and Safety Publications Series on Testing and Assessment. No. 184. [ENV/JM/MONO\(2013\)6](#), Paris, France.
- OECD, 2014a. Genotoxicity of Manufactured Nanomaterials : Report of the OECD Expert Meeting. OECD Environment, Health and Safety Publications Series on the Safety of Manufactured Nanomaterials. No. 43. [ENV/JM/MONO\(2014\)34](#), Paris, France.
- OECD, 2014b. Guidance document on an Integrated Approach on Testing and Assessment (IATA) for skin corrosion and irritation. Environment, Health and Safety Publications, Series on Testing and Assessment. No. 203. [ENV/JM/MONO\(2014\)19](#), Paris, France.
- OECD, 2015. Testing programme of manufactured nanomaterials – Dossiers and Endpoints. <http://www.oecd.org/chemicalsafety/nanosafety/dossiers-and-endpoints-testing-programme-manufactured-nanomaterials.htm>
- Ong, K.J. et al., 2014. Widespread nanoparticle-assay interference: implications for nanotoxicity testing." *PLoS One*, e90650.
- Oomen, A.G. et al., 2014. Concern-driven integrated approaches to nanomaterial testing and assessment – report of the NanoSafety Cluster Working Group 10. *Nanotoxicology*, 8(3), 334–348.
- Petersen, E.J. et al., 2014. Identification and avoidance of potential artifacts and misinterpretations in nanomaterial ecotoxicity measurements. *Environmental Science and Technology*, 48, 4226-4246.
- Roth, A. and Singer, T., 2014. The application of 3D cell models to support drug safety assessment: Opportunities and challenges. *Advanced Drug Delivery Reviews*, 69-70, 179–189.
- Rothen-Rutishauser, B.M., Kiama, S.G. and Gehr, P., 2005. A three-dimensional cellular model of the human respiratory tract to study the interaction with particles. *American Journal of Respiratory Cell and Molecular Biology*, 32(4), 281–289.
- Safe Work Australia, 2013. Development of an automated high-throughput screening procedure for nanomaterials genotoxicity assessment.
- Sandri, G. et al., 2010. Insulin-loaded nanoparticles based on n-trimethyl chitosan: *in vitro* (Caco-2 model) and *ex vivo* (excised rat jejunum, duodenum, and ileum) evaluation of penetration enhancement properties. *AAPS PharmSciTech*, 11(1), 362–371.
- Sauer, U.G. et al., 2014. Applicability of rat precision-cut lung slices in evaluating nanomaterial cytotoxicity, apoptosis, oxidative stress, and inflammation. *Toxicology and Applied Pharmacology*, 276(1), 1–20.
- SCCS, 2012. Guidance on the safety assessment of nanomaterials in cosmetics. SCCS/1484/12.
- Sharma, M. et al., 2016. Framework to evaluate exposure relevance and data needs for risk assessment of nanomaterials using *in vitro* testing strategies. *Risk Analysis*, DOI: 10.1111/risa.12581.

- Shatkin, J. A. and Ong, K. J. 2016. Alternative Testing Strategies for Nanomaterials: State of the Science and Considerations for Risk Analysis. *Risk Analysis*. doi:10.1111/risa.12642
- Shatkin, J.A., Ong, K.J., Beaudrie, C.B., Clippinger, A.J., Hendren, C.O., Haber, L.T., Hill, M., Holden, P., Kennedy, A.J., Kim, B., MacDonell, M., Powers, C.M., Sharma, M., Sheremeta, L., Stone, V., Sultan, Y., Turley, A., White, R.H. 2016. Advancing risk analysis for nanoscale materials: Report from an international workshop on the role of Alternative Testing Strategies. In press
- Stampfl, A. et al., 2011. Langendorff heart: A model system to study cardiovascular effects of engineered nanoparticles. *ACS Nano*, 5(7), 5345–5353.
- Stone, V, Johnston, HJ, Balharry, D., Gernand, JM, Gulumian, M. Approaches to develop alternative testing strategies to inform human health risk assessment of nanomaterials. *Risk Analysis*, 2016. *In press*
- Sønnegaard Poulsen, M. et al., 2013. Kinetics of silica nanoparticles in the human placenta. *Nanotoxicology*, 1–28.
- Tantra, R. et al., 2014. Nano(Q)SAR: Challenges, pitfalls and perspectives. *Nanotoxicology*, 1–7.
- Vecchio, G. et al., 2014. Lab-on-a-chip-based high-throughput screening of the genotoxicity of engineered nanomaterials. *Small*, 10(13), 2721–2734.
- Wang, A. et al., 2013. Characterization of physicochemical properties of nanomaterials and their immediate environments in high-throughput screening of nanomaterial biological activity. *Wiley Interdisciplinary Reviews: Nanomedicine and Nanobiotechnology*, 2013, 5, 430-448.
- Weir, A. et al., 2012. Titanium dioxide nanoparticles in food and personal care products. *Environmental Science and Technology*, 46(4), 2242–2250.
- Winkler, D.A. et al., 2013. Applying quantitative structure–activity relationship approaches to nanotoxicology: Current status and future potential. *Toxicology*, 313(1), 15–23.
- Wohlleben, W. et al., 2011. Artifacts by marker enzyme adsorption on nanomaterials in cytotoxicity assays with tissue cultures. *Journal of Physics: Conference Series*, 304, 012061.
- Wörle-Knirsch, J.M. et al., 2006. Oops they did it again! Carbon nanotubes hoax scientists in viability assays. *Nano Letters*, 6(6), 1261-1268.
- Xia, T. et al., 2013. Interlaboratory evaluation of *in vitro* cytotoxicity and inflammatory responses to engineered nanomaterials: The NIEHS Nano GO Consortium. *Environmental Health Perspectives*, 121(6), 683–690.
- Zhang, H. et al., 2012. Use of metal oxide nanoparticle band gap to develop a predictive paradigm for oxidative stress and acute pulmonary inflammation. *ACS Nano*, 6(5), 4349–4368.

ANNEX: AGENDA

Monday, September 15, 2014

12:30 pm	Welcome and Introductions: George Gray, George Washington University Jo Anne Shatkin, Society for Risk Analysis and Vireo Advisors, LLC
12:45 pm	The Context – Alternative Testing Strategies for Risk Analysis throughout the Product Life Cycle: Human Health, Ecosystem, Exposure <i>Presenter:</i> Maila Puolamaa, European Commission <ul style="list-style-type: none"> • 25 minute presentation + 10 min audience Q&A
1:05 pm	State of the Science: Applying a Multiple Models Approach to Inform Human Health Risk Assessment <i>Presenter:</i> Vicki Stone, Heriot Watt University <i>Respondents:</i> Myriam Hill, Health Canada; Monita Sharma, PETA International Science Consortium; Andrew Worth, Joint Research Centre; <ul style="list-style-type: none"> • 25 minute presentation + 20 minute discussion (3 respondents) • 15 min audience Q&A
2:05 pm	State of the Science: Applying a Multiple Models Approach to Inform Ecological Risk Assessment <i>Presenter:</i> Patricia Holden, University of California Center for Environmental Implications of Nanotechnology (UC CEIN); Bren School, UC Santa Barbara <i>Respondents:</i> Greg Goss, University of Alberta; Stacey Harper, Oregon State University; Yasir Sultan, Environment Canada <ul style="list-style-type: none"> • 25 minute presentation + 20 minute discussion (3 respondents) • 15 min audience Q&A
3:05 pm	<i>Coffee Break</i>
3:20 pm	State of the Science: Applying a Multiple Models Approach to Inform Exposure Assessment <i>Presenter:</i> Amy Clippinger, PETA International Science Consortium Ltd. <i>Respondents:</i> Shaun Clancy, Evonik; Steffi Friedrichs, Nanotechnology Industries Association; Treye Thomas, Consumer Product Safety Commission <ul style="list-style-type: none"> • 25 minute presentation + 20 minute discussion (3 respondents) • 15 min audience Q&A
4:20 pm	<i>Roundtable 1</i> – Using ATS for Screening Level Decisions in Human Health Risk Assessment <i>Panelists:</i> Maria Doa, United States Environmental Protection Agency; Mary Gulumian, South Africa National Institute for Occupational Health; Amy Wang, former US EPA <ul style="list-style-type: none"> • 25 minute panel discussion +10 minute Q&A

4:55 pm	<p><i>Roundtable 2 – Using ATS for Screening Level Decisions in Ecological Risk Assessment</i></p> <p><i>Panelists:</i> Christine Hendren, Duke University and Center for Environmental Implications of NanoTechnology; Alan Kennedy, UC Army Corps of Engineers Iseult Lynch, University of Birmingham</p> <ul style="list-style-type: none"> • 25 minute panel discussion + 10 minute Q&A
5:30 pm	<i>Poster Session / Reception to Follow</i>

Tuesday, September 16, 2014

8:30 am	Insights from Day 1
8:45 am	<p><i>Roundtable 3 – Approaches for Exposure Considerations in ATS</i></p> <p><i>Panelists:</i> Richard Canady, ILSI Risk Sciences Institute; Lynne Haber, Toxicology Excellence in Risk Assessment; Tom van Teunenbroek, Ministry Infrastructure and the Environment, The Netherlands</p> <ul style="list-style-type: none"> • 25 minute panel discussion + 10 minute Q&A
9:20 am	Overview of Breakout Sessions and Deliverables
9:50 am	Breakout Groups – Working Session #1
10:15 am	<i>Break</i>
10:30 AM	Breakout Groups – Working Session #1 cont'd
12:00 pm	<i>Lunch</i>
1:00 pm	Breakout Groups – Working Session #2
2:45 pm	<i>Coffee Break</i>
3:00 pm	Breakout Group Presentations and Plenary Discussion
4:10 pm	<p>Closing Session and Discussion of Next Steps</p> <ul style="list-style-type: none"> • <i>Timeline for products</i>
5:00 pm	<i>Adjourn</i>